

**IMPLEMENTASI *DATA MINING* UNTUK MEMPREDIKSI KELULUSAN
MAHASISWA FAKULTAS KOMUNIKASI DAN INFORMATIKA UMS
MENGUNAKAN METODE *NAIVE BAYES***

Makalah

Program Studi Teknik Informatika

Fakultas Komunikasi Dan Informatika



Diajukan Oleh :

Ahmad Fikri Mauriza

Yusuf Sulistyo Nugroho, S.T., M.Eng.

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA**

JANUARI 2014

HALAMAN PENGESAHAN

Publikasi Ilmiah dengan Judul

**IMPLEMENTASI *DATA MINING* UNTUK MEMPREDIKSI KELULUSAN
MAHASISWA FAKULTAS KOMUNIKASI DAN INFORMATIKA UMS
MENGUNAKAN METODE *NAIVE BAYES***

Yang dipersiapkan dan disusun oleh :

Ahmad Fikri Mauriza

L200100123

Telah Disetujui Pada:

Hari : Sabtu

Tanggal : 25 Januari 2019

Pembimbing,


Yusuf Sulisty Nugroho, S.T., M.Eng

Nik : 100.1197

Publikasi Ilmiah Ini Telah Diterima Sebagai Salah Satu Persyaratan

Untuk Memperoleh Gelar Sarjana

Tanggal.....

Mengetahui,

Ketua Program Studi

Teknik Informatika



Dr. Heru Supriyono, M.Sc.

Nik : 970



**UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI TEKNIK INFORMATIKA**

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id> Email: informatika@fki.ums.ac.id

SURAT KETERANGAN LULUS PLAGIASI

/A.3-II.3/INF-FKI/II/2014

Assalamu'alaikum Wr. Wb

Biro Skripsi Program Studi Teknik Informatika menerangkan bahwa :

Nama : AHMAD FIKRI MAURIZA
NIM : L200100123
Judul : IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI
KELULUSAN MAHASISWA FAKULTAS KOMUNIKASI DAN
INFORMATIKA UMS MENGGUNAKAN METODE NAÏVE
BAIYES
Program Studi : Teknik Informatika
Status : **Lulus**

Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Skripsi, dengan menggunakan aplikasi Turnitin.

Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb

Surakarta, 6 Februari 2014

Biro Skripsi
Teknik Informatika

Fauzan Nalsir, S.Kom

Turnitin Originality Report

IMPLEMENTASI DATA MINING UNTUK
MEMPREDIKSI KELULUSAN MAHASISWA
FAKULTAS KOMUNIKASI DAN
INFORMATIKA UMS MENGGUNAKAN
METODE NAIVE BAYES by Ahmad Fikri
Mauriza

Similarity Index

14%

Similarity by Source

Internet Sources	7%
Publications	1%
Student Papers	11%

From publikasi_maret_2014 (publikasi
maret 2014)

sources:

Processed on 04-Feb-2014 13:22 WIT
ID: 393366374
Word Count: 2514

1 8% match (student papers from 04-Feb-2014)
Class: publikasi maret 2014
Assignment:
Paper ID: [393366361](#)

2 1% match (Internet from 22-Oct-2010)

http://jas.eepis-its.edu/?page_id=23

3 1% match (student papers from 25-Jul-2013)

[Submitted to Universitas Muhammadiyah Surakarta on 2013-07-25](#)

4 1% match (Internet from 14-Nov-2008)

<http://www.ee.usyd.edu.au/~rafa/papers/08-alzoubi-koprinska-calvo.pdf>

5 1% match (Internet from 09-Nov-2012)

<http://magisterunlam.googlecode.com/files/Denny%20Setiawan%20C26209016.pdf>

6 1% match (Internet from 11-Apr-2011)

<http://eprints.undip.ac.id/view/subjects/QA75.html>

7 1% match (student papers from 04-Feb-2014)

Class: publikasi maret 2014

Assignment:

Paper ID: [393366363](#)

8 < 1% match (Internet from 03-Jul-2003)

<http://arxiv.org/abs/cond-mat/0307001>

9 < 1% match (Internet from 15-Sep-2010)

<http://www.scriptintermedia.com/view.php?id=5152&jenis=Indonesia>

IMPLEMENTASI DATA MINING UNTUK MEMPREDIKSI KELULUSAN MAHASISWA FAKULTAS KOMUNIKASI DAN INFORMATIKA UMS MENGGUNAKAN METODE NAIVE BAYES

Ahmad Fikri Mauriza, Yusuf Sulisty Nugroho
Teknik Informatika, Fakultas Komunikasi dan Informatika,
Universitas Muhammadiyah Surakarta
Email : Ahmadfikrimauriza@gmail.com

ABSTRACT

Faculty of Communication and Informatics in UMS has 2358 students and 388 graduated students. By the great numbers of new student each year, therefore, automatically it will create a huge number of data and the data stored in the database. Will increase significantly In order to make the big data to be a very useful information, the writer used data mining technique to predict the students length of study. It is to measure whether they can graduate less than or equal to 4 years or more based on the atributte used. Naive bayes method is used to analyze in the knowing of pattern and to predict the students graduation. The data needed is data from the graduated students as data training while the data from the active students is for testing data. Based on the result of prediction from 342 tested data sample, there are only 86 students who will graduate on time or only 25,15 % and the students who will be postponed in graduation is 256 students or it reaches 74,85 % from the whole data sample.

Keywords : Data Mining, Naive Bayes, Prediction

ABSTRAK

Fakultas komunikasi dan informatika universitas muhammadiyah surakarta memiliki jumlah 2358 mahasiswa dan memiliki lulusan 388 mahasiswa. Dengan bertambah pesatnya jumlah mahasiswa baru setiap tahunnya maka secara otomatis akan terciptanya data-data yang sangat banyak dan akan bertambah melimpahnya data yang tersimpan dalam database. Agar data yang awalnya sangat minim inoformasi tersebut dapat diubah menjadi suatu informasi yang sangat bermanfaat dengan menggunakan teknik data mining. Salah satunya dapat untuk memprediksi lama studi mahasiswa apakah dapat lulus kurang dari sama dengan 4 tahun atau lebih dari 4 tahun berdasarkan atribut-atribut yang digunakan. Metode naive bayes digunakan untuk menganalisis data dalam pengenalan pola dan memprediksi kelulusan mahasiswa. Data yang dibutuhkan adalah data-data mahasiswa yang sudah lulus sebagai data training sedangkan

data-data mahasiswa yang masih menempuh studi kuliah sebagai data uji. Berdasarkan hasil prediksi dari jumlah 342 data sampel yang diuji hanya 86 mahasiswa yang akan lulus tepat waktu atau hanya sekitar 25,15 %, sedangkan mahasiswa yang akan lulus terlambat berjumlah 256 mahasiswa mencapai 74,85% dari jumlah data sampel.

Kata Kunci : Data Mining, Naive Bayes, Prediksi

PENDAHULUAN

Universitas Muhammadiyah Surakarta merupakan salah satu universitas yang ada di Indonesia yang terdiri beberapa fakultas. Fakultas Komunikasi dan Informatika merupakan salah satunya yang berada di Universitas Muhammadiyah Surakarta. Sampai bulan Oktober 2013 kini fakultas ini sudah memiliki 2358 mahasiswa dengan jumlah lulusan 388 mahasiswa. Jumlah tersebut sudah tidak kecil lagi bagi fakultas yang sudah berusia 7 tahun ini, Fakultas tentu memiliki segudang data mengenai mahasiswanya meliputi: data induk mahasiswa dan data kelulusan mahasiswa.

Bertumpuknya data-data ini tentu belum digunakan secara optimal padahal tumpukan-tumpukan data tersebut dapat dijadikan suatu informasi yang sangat berharga. Untuk membantu dalam menemukan

informasi-informasi berharga tersebut diperlukan adanya teknik *data mining*. *Data mining* adalah teknik untuk menemukan dan mendeskripsikan pola-pola yang ada dalam data sebagai sebuah alat untuk membantu menjelaskan data tersebut dan membuat prakiraan dari data itu (Witten & Eibe Frank, 2005).

Dalam buku panduan akademik Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta tahun 2012 pada BAB 2 disebutkan bahwa “beban mahasiswa program pendidikan strata satu (S-1) beban studi harus ditempuh minimal 144 satuan kredit semester (SKS) dan maksimal 148 satuan kredit semester (SKS), dengan batas waktu kurang dari 8 semester dan waktu paling lama 14 semester”.

Tujuan dari penelitian ini yaitu memprediksi masa studi mahasiswa di Fakultas Komunikasi dan

Informatika Universitas
Muhammadiyah Surakarta
menggunakan metode *Naive Bayes*.

LANDASAN TEORI

1. Prediksi

Prediksi adalah memperkirakan sesuatu yang akan terjadi pada masa yang mendatang. Prediksi didapatkan melalui metode ilmiah maupun hanya subjektif belaka. Prediksi juga dapat digunakan dalam pengklasifikasian, tidak hanya untuk memprediksi *time series*, karena sifatnya yang bisa menghasilkan *class* berdasarkan berbagai atribut yang kita sediakan (Susanto, 2012).

2. Kelulusan Mahasiswa

Kelulusan mahasiswa tanda berakhirnya mahasiswa dalam menyelesaikan pendidikan pada jenjang sarjana. Kelulusan juga merupakan hasil akhir pencapaian yang membanggakan dalam menempuh suatu pendidikan pada jenjang tertentu. Beban mahasiswa program pendidikan strata satu (S-1) beban studi harus ditempuh minimal 144 satuan kredit semester (SKS) dan maksimal 148 satuan kredit semester (SKS), dengan batas waktu kurang

dari 8 semester dan waktu paling lama 14 semester. Mahasiswa dinyatakan lulus apabila ketentuan diatas dilaksanakan dan tidak ada tanggungan administrasi di biro adminitrasi akademik (BAA) (Anonim, 2012).

3. Naive Bayes

Naive Bayes merupakan teknik prediksi berbasis probabilistik sederhana yang berdasar pada penerapan *Teorema Bayes* (atau aturan Bayes) dengan asumsi *independensi* (ketidaktergantungan) yang kuat. (Prasetyo, 2012)

METODOLOGI PENELITIAN

a. Penentuan Atribut

Dalam memprediksi kelulusan mahasiswa maka diperlukan data-data seluruh mahasiswa baik yang sudah lulus maupun yang belum lulus. Semua data-data diperoleh dari Biro Administrasi Akademik (BAA) dan Laboratorium Fakultas Komunikasi dan Informatika. Data yang digunakan sebagai data training yaitu mahasiswa yang sudah lulus dari angkatan 2007 sampai angkatan 2009, sedangkan data yang digunakan sebagai data uji yaitu mahasiswa yang

masih aktif mengikuti perkuliahan diambil sampel secara acak. Atribut-atribut yang akan digunakan yaitu:

Tabel 1 Daftar Atribut

Atribut	Isi dalam Atribut-Atribut
Jurusan asal sekolah	IPA, IPS dan Lain
Gender	Pria dan Wanita
Daerah asal mahasiswa	Surakarta dan Luar Surakarta
Asal sekolah	Surakarta dan Luar Surakarta
Rata SKS (satuan kredit semester)	$SKS \leq 18$ dan $SKS > 18$
Rata MK (Mata kuliah)	$MK \leq 7$ dan $MK > 7$
Asisten Lab	Ya dan Tidak
Lama studi	Tepat (lulus ≤ 4 tahun) dan Terlambat (lulus > 4 tahun)

b. Penentuan sampel

Untuk menentukan jumlah sampel yang akan digunakan dalam menggambarkan dari jumlah populasi menggunakan bantuan metode slovin dengan nilai maksimal $e = 5\%$ (Umar, 2004).

$$n = \frac{N}{1 + N e^2} \quad (1)$$

Bahwa :

n = jumlah sampel

N = jumlah populasi

e = nilai toleransi

c. Diubah dalam format Arff

Data-data yang telah diolah dalam *microsoft excel* kemudian diubah ke dalam format Arff (*Attribute Relation File Format*) supaya dapat diolah dalam aplikasi *weka*.

d. Penggunaan Metode

Metode yang digunakan dalam penelitian ini yaitu *naive bayes* untuk memprediksi masa studi mahasiswa Fakultas Komunikasi dan Informatika. Pemilihan metode ini dikarenakan relatif mudah digunakan karena tidak ada perkalian matrik atau optimasi numerik, lebih efisien apabila digunakan untuk memprediksi dalam jumlah yang sangat besar, dan memiliki tingkat keakurasian yang relatif tinggi dalam hasil prediksi.

Metode *naive bayes* juga sering disebut dengan algoritma HMAP (*Hypothesis Maximum Apriori*

Probability) yang merupakan penyederhanaan dari metode *bayes*. Metode ini menyatakan hipotesa dari penghitungan menggunakan probabilitas berdasarkan kondisi *prior* (Prasetyo, 2012).

$$\begin{aligned} P(S|X) &= \operatorname{argmax}_{x \in X} \frac{P(Y|X) P(X)}{P(X)} \\ &= \operatorname{argmax}_{x \in X} P(Y|X) P(X) \end{aligned} \quad (2)$$

HASIL DAN PEMBAHASAN

a. Pengambilan sampel

Apabila diketahui mahasiswa Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta memiliki jumlah 2358 mahasiswa dan memiliki toleransi ketidakteelitian 5%. Maka jumlah sampel yang diambil yaitu:

$$n = \frac{N}{1 + N e^2}$$

$$n = 2358 / 1 + 2358 \times (0,05)^2$$

$$n = 2358 / 1 + 2358 \times 0,0025$$

$$n = 2358 / 1 + 5,895$$

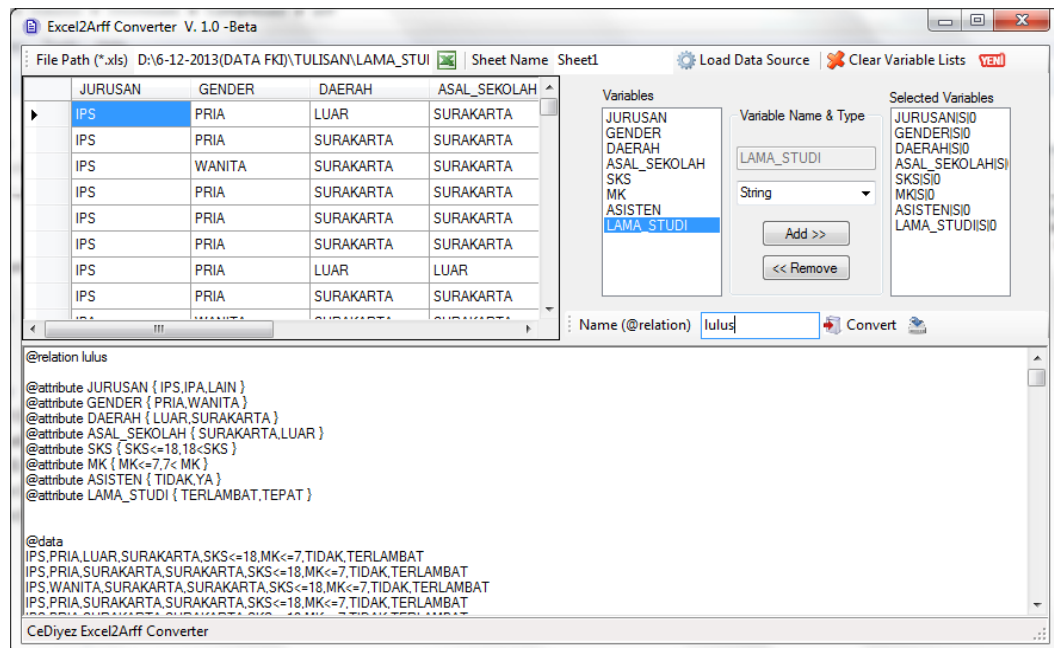
$$n = 2358 / 6,895$$

$$n = 341,9869 \text{ mahasiswa}$$

Jadi dibulatkan menjadi 342 mahasiswa yang digunakan bahan sampling untuk data uji.

b. Pengubahan format

Pengubahan format dari format .xls diubah menjadi .arff diperlukan aplikasi pembantu yaitu *Excel2Arff Converter*. Agar dapat dibuka dan diolah dalam aplikasi *weka*. Gambar pemrosesan konverter dapat dilihat pada Gambar.1.



Gambar 1. Proses konverter

c. Perhitungan *Naive Bayes*

Sebagai contoh di ambikan satu data uji yang memiliki ciri sebagai berikut: IPS, Pria, memiliki alamat luar surakarta, asal sekolah luar surakarta, $SKS \leq 18$, $MK \leq 7$, tidak anggota asisten. Apakah mahasiswa tersebut akan lulus tepat waktu atau terlambat?

Penghitungan data test berdasarkan data training:

Asumsi :

Y = Lama studi

X_1 = Jurusan SMA/SMK

X_2 = Gender

X_3 = Daerah / Alamat

X_4 = Asal sekolah

X_5 = Rata-rata SKS

X_6 = Rata-rata mata kuliah

X_7 = Asisten

Fakta menunjukkan :

$$P(Y = \text{TEPAT}) = 73/341 = 0,214076$$

$$P(Y = \text{TERLAMBAT}) = 268/341 = 0,785924$$

Fakta :

$$P(X_1 = \text{IPS} | Y = \text{TEPAT}) = 22/73 = 0,30137$$

$$P(X_1 = \text{IPS} | Y = \text{TERLAMBAT}) = 127/268 = 0,473881$$

$$P(X_2 = \text{PRIA} | Y = \text{TEPAT}) = 46/73 = 0,630137$$

$$P(X_2 = \text{PRIA} | Y = \text{TERLAMBAT}) = 165/268 = 0,615672$$

$$P(X_3=\text{LUAR} | Y=\text{TEPAT}) = 14/73 \\ = 0,191781$$

$$P(X_3=\text{LUAR} | Y=\text{TERLAMBAT}) = \\ 78/268 = 0,291045$$

$$P(X_4=\text{LUAR} | Y=\text{TEPAT}) = 14/73 \\ = 0,191781$$

$$P(X_4=\text{LUAR} | Y=\text{TERLAMBAT}) = \\ 70/268 = 0,261194$$

$$P(X_5=\text{SKS} \leq 18 | Y=\text{TEPAT}) = 4/73 \\ = 0,054795$$

$$P(X_5=\text{SKS} \leq 18 | Y=\text{TERLAMBAT}) = \\ 166/268 = 0,619403$$

$$P(X_6=\text{MK} \leq 7 | Y=\text{TEPAT}) = 0/73 = 0$$

$$P(X_6=\text{MK} \leq 7 | Y=\text{TERLAMBAT}) = \\ 141/268 = 0,526119$$

$$P(X_7=\text{TIDAK} | Y=\text{TEPAT}) = 54/73 \\ = 0,739726$$

$$P(X_7=\text{TIDAK} | Y=\text{TERLAMBAT}) = \\ 248/268 = 0,925373$$

HMAP dari keadaan ini dapat dihitung dengan :

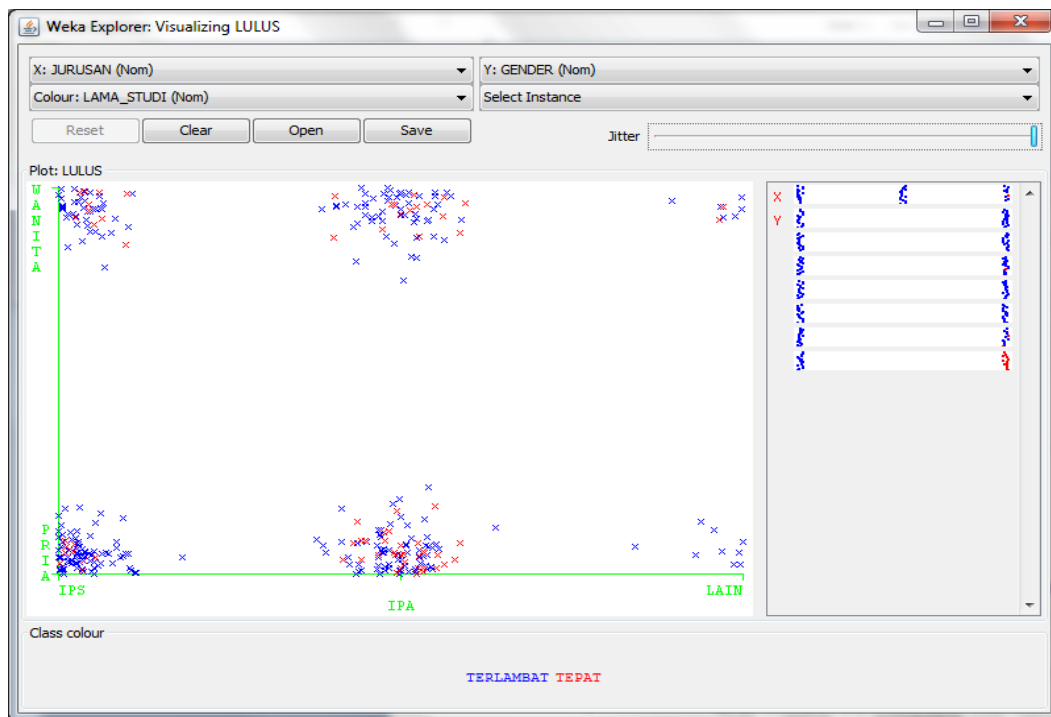
$$P(X_1=\text{IPS}, X_2=\text{PRIA}, X_3=\text{LUAR}, \\ X_4=\text{LUAR}, X_5=\text{SKS} \leq 18, \\ X_6=\text{MK} \leq 7, X_7=\text{TIDAK} | Y \\ =\text{TEPAT}) \\ = 22/73 * 46/73 * 14/73 * 14/73 \\ * 4/73 * 0/73 * 54/72 * 73/341 \\ = 0$$

$$P(X_1=\text{IPS}, X_2=\text{PRIA}, \\ X_3=\text{LUAR}, X_4=\text{LUAR}, \\ X_5=\text{SKS} \leq 18, X_6=\text{MK} \leq 7, \\ X_7=\text{TIDAK} | Y \\ =\text{TERLAMBAT}) \\ = 127/268 * 165/268 * 78/268 * \\ 70/268 * 166/268 * 141/268 * \\ 248/268 * 268/341 \\ = 0,005256$$

❖ **KEPUTUSAN LAMA STUDI =
TERLAMBAT.**

d. Pengolahan dalam Weka

Data training yang sudah diubah dalam format .arff kemudian akan diolah menggunakan aplikasi weka guna mengetahui pengklasifikasian mahasiswa yang lulus dengan tepat waktu atau terlambat. Dataset yang akan menjadi data trining yaitu 341 record dan 8 atribut. Setelah diproses weka akan menampilkan grafik yang menunjukkan hasil pengklasifikasian sesuai atribut masing-masing dan pengaruh atribut lainnya.

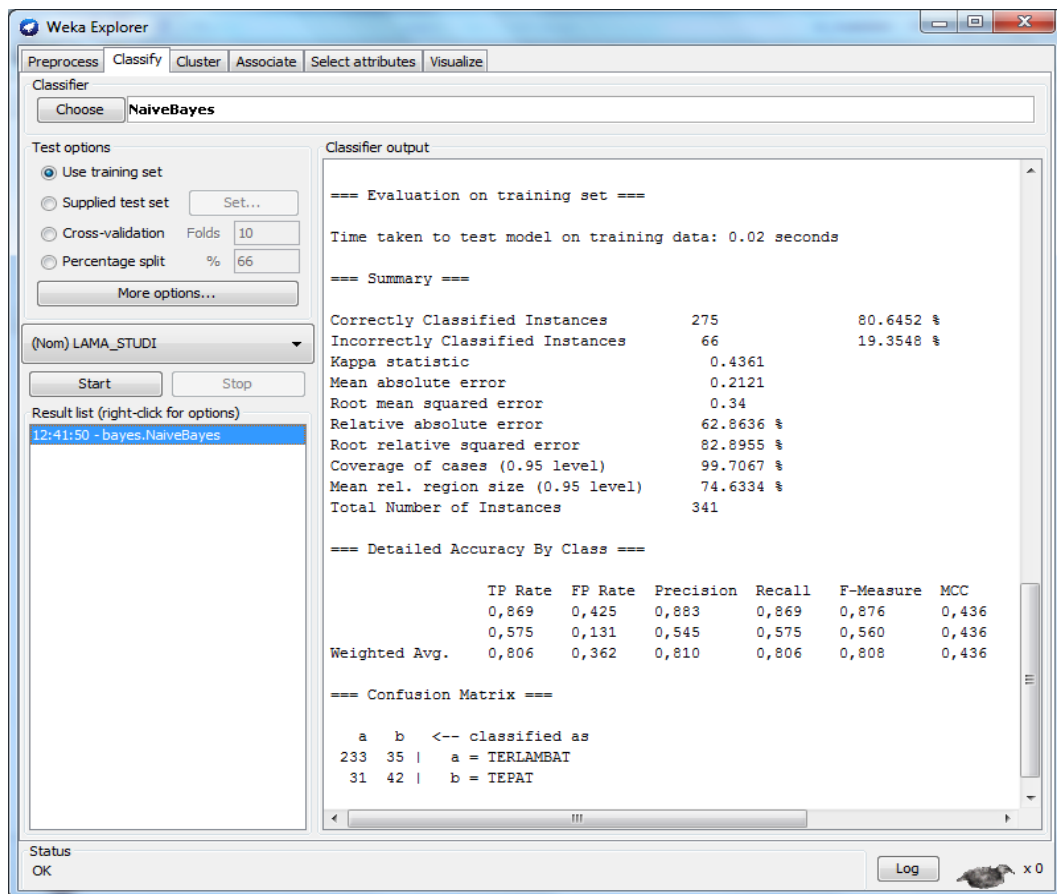


Gambar 2. Visualize jurusan dan gender.

Gambar 2 memiliki dua sumbu yaitu sumbu X dan sumbu Y, sumbu X yaitu jurusan sedangkan sumbu Y yaitu gender. Pada titik merah menunjukkan mahasiswa yang lulus dengan tepat waktu sedangkan titik yang berwarna biru menunjukkan mahasiswa yang telah lulus terlambat. Pada Gambar 2 dapat dilihat bahwa mahasiswa yang lulus tepat waktu lebih banyak berasal dari jurusan IPA dan berjenis kelamin Pria.

Data training tersebut akan diproses menggunakan metode

classifier naive bayes dengan hasil luaran status lama studi. Dari data tersebut akan membentuk pola sehingga pola-pola yang sudah terbentuk dapat digunakan untuk memprediksi lama studi untuk mahasiswa yang belum lulus.



Gambar 3. Evaluasi naive bayes

Data yang digunakan dalam training berjumlah 341 record didapatkan hasil seperti Gambar 3, sehingga dapat dilihat hasil presentase akurasiya yaitu :

$$\text{Persentase Akurasi} = \frac{\text{Banyak prediksi yang benar}}{\text{Total banyaknya data}} \times 100\%$$

Persentase Akurasi

$$\begin{aligned}
 &= (233+42) / (233+35+31+42) * 100\% \\
 &= 80,6452 \%
 \end{aligned}$$

Tahap berikutnya yaitu memprediksi data uji menggunakan data training yang sudah ada. Pada *Test option* pilih *Supplied test set* kemudian klik set untuk memasukan data uji kemudian diproses dan disimpan dalam format Arff. Sehingga hasil prediksi yang telah dilakukan oleh *weka* dapat diketahui melalui *ArffViewer*.

ARFF-Viewer - D:\6-12-2013(DATA FKJ)\TULISAN\LAMA_STUDI\HASIL_NAIVE.arff

File Edit View

HASIL_NAIVE.arff

Relation: LULUS_predicted

No.	1: JURUSAN Nominal	2: GENDER Nominal	3: DAERAH Nominal	4: ASAL_SEKOLAH Nominal	5: SKS Nominal	6: MK Nominal	7: ASISTEN Nominal	8: prediction margin Numeric	9: predicted LAMA_STUDI Nominal	10: L
1	IPS	PRIA	LUAR	LUAR	SKS(=18	MK(=7	TIDAK	0.99955	TERLAMBAT	
2	IPA	PRIA	LUAR	LUAR	SKS(=18	MK(=7	TIDAK	0.999043	TERLAMBAT	
3	IPS	PRIA	LUAR	LUAR	18(SKS	MK(=7	TIDAK	0.990086	TERLAMBAT	
4	LAIN	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.99894	TERLAMBAT	
5	IPS	WANITA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.999078	TERLAMBAT	
6	IPS	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.998984	TERLAMBAT	
7	IPS	WANITA	LUAR	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.99943	TERLAMBAT	
8	IPS	WANITA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.999078	TERLAMBAT	
9	IPS	PRIA	LUAR	LUAR	SKS(=18	MK(=7	YA	0.998244	TERLAMBAT	
10	IPA	PRIA	LUAR	LUAR	SKS(=18	MK(=7	TIDAK	0.999043	TERLAMBAT	
11	IPS	WANITA	LUAR	LUAR	SKS(=18	MK(=7	TIDAK	0.999592	TERLAMBAT	
12	IPS	PRIA	LUAR	SURAKARTA	18(SKS	MK(=7	TIDAK	0.986184	TERLAMBAT	
13	IPS	PRIA	SURAKARTA	SURAKARTA	18(SKS	7(MK	TIDAK	0.050057	TERLAMBAT	
14	IPS	WANITA	SURAKARTA	LUAR	SKS(=18	MK(=7	TIDAK	0.999339	TERLAMBAT	
15	IPS	WANITA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.999078	TERLAMBAT	
16	IPS	WANITA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.999078	TERLAMBAT	
17	IPS	WANITA	SURAKARTA	SURAKARTA	18(SKS	7(MK	TIDAK	0.097925	TERLAMBAT	
18	LAIN	WANITA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.999038	TERLAMBAT	
19	IPS	WANITA	SURAKARTA	LUAR	18(SKS	MK(=7	TIDAK	0.985461	TERLAMBAT	
20	IPA	WANITA	SURAKARTA	SURAKARTA	18(SKS	7(MK	TIDAK	-0.272834	TERPAT	
21	IPS	WANITA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.999078	TERLAMBAT	
22	IPS	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.998984	TERLAMBAT	
23	IPS	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.998984	TERLAMBAT	
24	IPS	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.998984	TERLAMBAT	
25	IPS	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	TIDAK	0.998984	TERLAMBAT	
26	IPS	PRIA	SURAKARTA	SURAKARTA	18(SKS	MK(=7	TIDAK	0.977735	TERLAMBAT	
27	IPA	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	YA	0.991572	TERLAMBAT	
28	IPS	PRIA	SURAKARTA	SURAKARTA	SKS(=18	MK(=7	YA	0.996035	TERLAMBAT	

Gambar 4. Hasil Prediksi Data Uji

Hasil pengklasifikasian mahasiswa yang lulus tepat waktu yaitu lulus kurang sama dengan 4 tahun didominasi berasal dari jurusan asal IPA, Jenis kelamin antara pria lebih baik daripada wanita. Untuk daerah asal mahasiswa dan sekolah asal SMA/SMK didominasi oleh Karesidenan Surakarta, dengan jumlah rata-rata pengambilan SKS lebih 18 SKS setiap semesternya dan rata-rata pengambilan jumlah mata kuliah lebih dari 7 mata kuliah. Asisten laboratorium ternyata tidak terlalu banyak berpengaruh terhadap ketepatan waktu mahasiswa lulus yaitu kurang atau sama dengan 4 tahun.

e. Membandingkan dengan metode lain

Untuk mengetahui tingkat keakurasian metode *naive bayes* dengan metode lain perlu adanya perbandingan antar metode satu dengan lainnya. Penulis ingin mengetahui lebih tinggi mana tingkat keakurasianya antara metode *naive bayes*, dengan metode *decicion tree* (J48) dan metode *OneR* dengan data yang digunakan sama.

Tabel 2 Perbandingan dari naive bayes, J48 dan OneR

	Naive bayes	J48	OneR
Akurasi	80.645 2 %	85,63 05 %	78.59 24 %
Waktu proses	0,01 s	0,05 s	0,01 s
Data benar	275 data	292 data	268 data
Data kurang sesuai	66 data	49 data	73 data

Dari perbandingan pada tabel 2 dapat disimpulkan bahwa metode J48 lebih baik dari pada metode *naive bayes* dan *OneR* dalam permasalahan ini. Sedangkan *naive bayes* lebih bagus dari pada *OneR* karena memiliki keakurasian hasil prediksi lebih baik. Metode J48 lebih unggul dalam penelitian ini karena data yang digunakan jumlahnya kurang banyak sehingga metode J48 lebih cepat dan mudah dalam pengklasifikasian. Metode J48 (*decision tree*) merupakan memang algoritma yang digunakan untuk pengelompokan *class*. Namun untuk waktu pemrosesan (running time) yang paling cepat yaitu

Naive Bayes dari ketiga algoritma tersebut.

KESIMPULAN

Dari jumlah sampel 342 mahasiswa yang digunakan data uji menggunakan metode *naive bayes* ternyata didapatkan hasil bahwa mahasiswa yang akan lulus tepat waktu hanya berjumlah 86 mahasiswa atau hanya sekitar 25,15 % dari jumlah sampel sedangkan mahasiswa yang akan lulus terlambat berjumlah 256 mahasiswa mencapai 74,85%.

SARAN

1. Bagi peneliti-peneliti selanjutnya yang akan melakukan penelitian hampir serupa dengan penelitian ini.
 - a) Mencoba menggunakan aplikasi selain *Weka* dalam analisa data mining.
 - b) Mencoba lebih banyak *record* dan *attribute* dalam pemrosesan data mining.
 - c) Diperlukan tingkat ketelitian yang tinggi dan pembersihan data yang sempurna supaya tidak terjadinya *noise*.

- d) Data perlu menyesuaikan dengan kurikulum yang terbaru.
 - e) Dibuatkan grafik jumlah lulusan setiap tahunnya agar mengetahui ada kenaikan ataukah penurunan.
2. Bagi Fakultas Komunikasi dan Informatika Universitas Muhammadiyah Surakarta berdasarkan hasil penelitian.
- a) Melihat hasil kesimpulan pada penelitian ini sungguh memprihatinkan karena dari jumlah sampel yang diprediksi hanya sekitar 25,15 % saja yang akan lulus tepat waktu sedangkan mahasiswa yang akan lulus terlambat mencapai 74,85%. Sehingga pihak fakultas perlu adanya strategi atau solusi dalam peningkatan prestasi mahasiswanya sehingga mahasiswa dapat lulus tepat waktu dan tentunya dengan hasil yang memuaskan. Sehingga dapat membantu fakultas dalam meningkatkan nilai akreditasi.
 - b) Berdasarkan kesimpulan mahasiswa yang dapat lulus tepat waktu yang paling menonjol yaitu dipengaruhi oleh pengambilan jumlah rata-rata SKS lebih dari 18 SKS dan lebih dari 7 mata kuliah setiap semesternya, mahasiswa yang paling banyak mengambil jumlah SKS dan Mata kuliah yaitu berasal dari jurusan IPA berjenis kelamin Pria berasal dari daerah Karesidenan Surakarta. Jika fakultas ingin memiliki mahasiswanya dapat lulus tepat waktu, maka mahasiswa baru yang akan masuk lebih diutamakan berasal dari Jurusan IPA, Pria , asal daerah dan sekolah berasal dari Karesidenan Surakarta.

DAFTAR PUSTAKA

- Anonim.2012 .*Panduan akademik fakultas komunikasi dan informatika UMS 2012/2013*. Surakarta : Universitas Muhammadiyah Surakarta.
- Handayanto, Rahmadya Trias. 2010. *Naive bayes*. From website <<http://rahmadya.com>>. Diakses pada tanggal 30 oktober 2013.
- Huda, Nuqson Masykur. 2010. *Aplikasi Data Mining Untuk Menampilkan Informasi Tingkat Kelulusan Mahasiswa*, Skripsi. Semarang: Fakultas MIPA Universitas Diponegoro.
- Prasetyo, Eko. 2012. *Data Mining konsep dan aplikasi menggunakan matlab*. Yogyakarta: Andi.
- Sundika, Putu. *Comparisson Of One-R, Decision Tree & Naive Bayes*. From website <http://www.slideshare.net/putusundika/oner-bayes-decisiontree>. Diakses pada tanggal 19 Desember 2013.
- Umar, Husein. 2004. *Metode Penelitian Untuk Skripsi Dan Tesis Bisnis*. Cetakan ke-6. Jakarta: PT Raja Grafindo Persada.
- Witten, I. H and Frank, E. 2005. *Data Mining : Practical Machine Learning Tools and Techniques Second Edition*. San Francisco: Elsevier.

BIODATA PENULIS

Nama : Ahmad Fikri Mauriza
Tempat / Tanggal Lahir : Jepara, 30 Maret 1992
Jenis Kelamin : Pria
Agama : Islam
Jurusan : Teknik Informatika
Peguruan Tinggi : Universitas Muhammadiyah Surakarta
Alamat : Jl. A. Yani Tromol Pos I Pabelan, Kartasura
Telp./ Fax : (0271)717417, 719483 / (0271) 714448
Alamat Rumah : Blingoh RT 12/ III Donorojo, Jepara
No. HP : +6285641788799
Alamat e-mail : Ahmadfikrimauriza@gmail.com